AWARD NUMBER: W81XWH-16-1-0107

TITLE: Identification of Immunogenic Targets for Lung Cancer Vaccines

PRINCIPAL INVESTIGATOR: Laura Riolobos

CONTRACTING ORGANIZATION: University of Washington
Seattle, WA 98195

REPORT DATE: September 2017

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
September 2017	Annual	1 Sept 2016 – 31 Aug 2017

4. TITLE AND SUBTITLE

Identification of Immunogenic Targets for Lung Cancer Vaccines

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0107

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Laura Riolobos

E-Mail: lriolobo@uw.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Washington
4333 Brooklyn Avenue
Seattle, WA 98195

8. PERFORMING ORGANIZATION REPORT

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The goal of this project is to develop a vaccine for lung cancer prevention in current or past smokers by identifying immunogenic proteins in lung cancer that are able to induce a potent inflammatory Th1 response. Proteins overexpressed in tumors can be immunogenic. We have used quantitative proteomic analysis to identify proteins overexpressed in non-small cell lung cancer cell lines compared with normal lung epithelial cells. Candidate antigens were investigated by siRNA screening to identify those genes with a function in maintaining cell tumor growth. If a gene is required for tumor cell proliferation, knocking down the gene by siRNA should decrease cell survival and proliferation. We identified 14 candidates that are overexpressed in lung cancer and necessary for tumor cell survival. We have prioritized those proteins which have been previously described to play a role in lung cancer invasion, proliferation, metastasis or survival. We selected 5 of those proteins to move forward. We have done in silico screening for HLA-DR epitope prediction and we have been able to identify epitopes with a high probability of been presented by multiples HLA-DR alleles. Therefore, during this period of the award, we have identified possible antigens and epitopes to develop a multi-antigen vaccine to prevent lung cancer.

15. SUBJECT TERMS Identification of antigens and epitopes for a lung cancer vaccine					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified	Unclassified	19	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

1. INTRODUCTION:

The goal of this project is to develop a vaccine for lung cancer prevention in current or past smokers by identifying immunogenic proteins in lung cancer that are able to induce a potent inflammatory Th1 response. Lung cancer has one of the highest mutation rates of all types of cancer, but driver mutations that could be targeted for a vaccine for lung cancer prevention are unknown. Gene expression profiling of bronchial biopsy specimens from smokers have shown that changes in gene expression arise in histologically normal epithelia and can be used to discern people with and without lung cancer.[1] These early changes in protein expression can be identified by auto-antibodies specific to tumor-associated antigens and can be detected in pre-neoplasic lesions.[2, 3] Many of these immunogenic proteins are aberrantly upregulated, but not mutated. Our group has previously shown that overexpressed proteins in tumors can be immunogenic.[4] Type I (Th1) cellular immunity is needed for cancer eradication.[5] Type I immunity enhances cross-priming of tumor antigens by activating local antigen presenting cells (APC) to more efficiently present immunogenic proteins to T-cells. Studies by our group have demonstrated that those class II epitopes, derived from self-proteins and most likely to induce a Th response, are epitopes that bind with high affinity across multiple HLA-DR alleles. We utilize web based algorithms for HLA-DR epitope prediction to identify peptides with a likelihood of high affinity promiscuous binding. We have developed ELISPOT based screening of those predicted epitopes focused on assessing IFN-g (Th1) and IL-10 (Th2) responses. Our rationale is to exclude epitopes from a cancer vaccine that may elicit a predominant Th2 response as those peptides would stimulate T-cells that could prevent the optimal function and proliferation of Th1 and CD8+ cytotoxic T cell (CTL) necessary for cancer eradication. In this project, we will use proteomic approaches to identify candidate proteins which are overexpressed in lung cancer cell lines that play a role in the tumor cell survival. Then, we will screen those antigens for selective Th1 inducing epitopes. Our expectation is that immunity against identified antigens, induced by Th1 epitopes, will prevent or inhibit the development of malignancy.

Specific Aims: (1) Identify proteins that are overexpressed in human lung cancer cell lines and determine their functional relevance to lung cancer progression and their immunogenicity; (2) Determine epitopes derived from the identified antigens that will preferentially elicit a Th1 CD4+ T cell response; and (3) Evaluate the antigen specific Th1 vaccines for prophylactic efficacy in mouse models of lung cancer.

2. KEYWORDS:

Non-small cell lung cancer, immunogenic antigens, preventive multi-antigen cancer vaccine, type I immunity

3. OVERALL PROJECT SUMMARY:

TASK 1. Identify proteins that are overexpressed in human lung cancer cell lines and determine their functional relevance to lung cancer progression and their immunogenicity.

Subtask 1.1. Identification of proteins overexpressed in human lung cancer cell lines by quantitative mass-spectroscopy.

1.1.A. Quantitative mass-spectroscopy of lung cancer cell lines (squamous cell carcinoma-SCC- and adenocarcinoma- AC) and normal epithelial lung cell lines. Selection of antigens that were overexpressed more than 1.2 fold in the cancer versus the normal cells.

The iTRAQ 8-plex (Sigma-Aldrich) allows multiplexing up to eight different samples in a single quantitative mass-spectrometry (LC/MS) experiment for protein quantification. Each sample is labelled with a stable isotope molecule (tag) that allowed identification and relative quantification between samples. The following cell lines were included for LC/MS: Nuli-1, BEAS-2B, HCC1588, HCC1313, NCI-H520, NCI-H522 and HCC461. Details of the type of cell (normal, squamous-SCC- or adenocarcinoma- AC), and stage of cancer when cells was derived from patients can be found in the Table 1. For each sample, 100µg of total protein was denature with 2% SDS, and reduced with 1% TCEP for one hour at 60C. After that, Cysteines were blocked with MMTS 200mM for 10 min at room temperature. In order to eliminate contaminants that interfere with following steps, each sample was precipitated with 6 volumes of Acetone at -20C for 4h, and dissolved in 50mM TEAB (triethyl ammonium bicarbonate buffer). Then, samples were treated with trypsin (10µg/sample) overnight at 37C. iTRAQ labeling was performed at room temperature for 2h. Nuli1 normal cell extract was labelled twice with 2 different tags (113 and 121) in order to asses our intra-experiment variability. The rest of the samples were labeled as follows: BEAS-2B 114-tag, HCC1588 115-tag, HCC1313 116-tag, H520 119-tag, H461 118-tag and H522 117-tag. After labeling, all samples were pooled and then clean up to eliminate the excess of unbound iTRAQ reagent using a cation-exchange chromatography (C18 columns, Pierce) following the recommendations of the vendor.

Sample was analyzed in LC/MS by Dr. Priska von Haller and preliminary data analysis and peptide identification was done in collaboration with Dr. Jimmy Eng, both of them from the UW Proteomic Core. All samples were analyzed on an Orbitrap Fusion Lumos mass spectrometer ThermoFisherScientific) equipped with a nano-Acquity UPLC system (Waters) and in house developed nano spray ionization source. Samples (2 µl at 1.5 µg/µl) were loaded from the autosampler onto a 100 µm ID Integrafrit trap (NewObjective) packed with Dr. Maisch Reprosil-Pur C18-AQ 120 Å 5 µm material to a bed length of 3 cm at a flow rate of 2 µL/min. After loading and desalting for 10 min with 0.1% formic acid plus 2% acetonitrile (LCMS grade from Fisher), the trap was brought in-line with a pulled fused-silica capillary tip (75-µm i.d.) packed with 35 cm of Dr. Maisch Reprosil-Pur C18-AQ 120 Å 5 µm. Peptides were separated using a linear gradient, from 5-30% solvent B (LCMS grade 0.1 % formic acid it acetonitrile (Fisher)) in 90 min at a flow rate of 300 nL/min. Peptides were detected using a data-dependant (DDA) method. Survey scans of peptide precursors were performed in the orbitrap mass analyzer from 375 to 1575 m/z at 120K resolution (at 200 m/z) with a 7e5 ion count target and a maximum injection time of 50 ms. The instrument was set to run in top speed mode with 3 sec cycle for the survey and the MS/MS scans. After a survey scan, tandem MS was performed on the most abundant precursors exhibiting a charge state from 2 to 5 of greater than 2e4 intensity by isolating them in the quadrupole with an isolation width of 1.2 m/z. Higher-energy collisional dissociation (HCD) fragmentation was

applied with a normalized collision energy of 32 %. Resulting fragments were detected in the orbitrap mass analyzer at 50K resolution (at 200 m/z) with a 1e5 ion count target and a maximum injection time of 100 ms. The dynamic exclusion was set to 30 sec with a 10 ppm mass tolerance around the precursor and its isotopes. Monoisotopic precursor selection was enabled.

Mass spectra generated in the analysis was used to identify peptide sequences with the Comet search engine (version 2016.01.2) by searching the experimental tandem mass spectra against the human sequence database UniProtKB. The human UniProtKB database was downloaded from UniProt on Jan 23, 2017 and appended with common contaminant sequences (cRAP contaminant proteins from the Global Proteome Machine). There were a total of 92,976 sequence entries in the database file (92,928 human sequences and 48 contaminant sequences). The Trans-Proteomics Pipeline suite of tools (version 5.0.0) was subsequently applied for downstream analysis including: calculating peptide identification confidence with PeptideProphet, inferring proteins with ProteinProphet, and calculating quantitative abundances of the Tandem Mass Tag (TMT) stable isotope labeled sample with Libra.

We were able to identify a total of 14,219 different peptides, which correspond to 2,875 proteins. The method to identify proteins overexpressed in the tumor versus the normal cell line is shown in Figure 1. Briefly, we looked for those proteins that were identified with (a) 100% confidence or (b) 95% confidence AND at least 3 peptides identified for that specific protein. Then, for each cell line, we normalized with the expression of the protein actin-beta (ACTB) to compensate for possible loading differences between samples. Finally, we compared for each cell line and protein, the relative expression to the normal cell line Nuli-1. We considered that a protein is overexpressed if: [expression in cell line/ expression in normal Nuli-1] > 1.5. We selected those proteins overexpressed in at least 3 or more of the lung cancer cell lines screened: 39 proteins from the 100% confidence group and 138 from the 95% confidence. There was an overlapping of 23 proteins between the 100% and the 95% confidence group. Therefore, a total of 154 antigens were selected for siRNA screening.

Subtask 1.2. Determine candidate overexpressed proteins that have functional relevance in proliferation and apoptosis by siRNA screening.

1.2.A. High-throughput screening of candidate genes for cell survival using siRNA knockdown

The 154 genes identified in the proteomic analysis were evaluated by siRNA screening in collaboration with Quellos facility at the UW. One normal lung epithelial cell line (Nuli-1) and 4 non-small cell (NSCLC) lung cancer cell lines, 2 squamous cell carcinoma (HCC1588 and H520) and 2 adenocarcinoma (H522 and H461), were included in the analysis. Feasibility experiments were conducted in advance to determine the number of cells and the dose of transfection reagent to obtain an optimal separation between the viability curves for a negative control siRNA and the highly toxic siRNA targeting the kinesis motor protein Kiff11. Based on our optimization experiments, conditions for the experiments were 500 cells per well, 12.5 µl/ml of RNAiMax transfection mix, and analysis at 96 hours post transfection. We looked for those siRNA that decreased the viability of the cancer cell lines but not the normal cell line. We considered for further analysis those genes for which [(mean of viability tumor cell line) / (mean of viability in the normal cell line)] <0.75 with a p-value <0.1 for any of the cell lines assayed. We identified 21 genes that met our criteria (Table 2).

1.2.B. For candidate genes with decreased survival in lung cancer cell lines, test siRNA induction of apoptosis with Caspase 3/7 and DNA damage assays

We analyzed if the decrease in viability correlated with increased apoptosis by Caspase 3/7. As before, one normal cell line (Nuli-1) and 4 NSCLC cell lines (HCC1588, H522, H520 and H461) were included in the analysis. We transfected cells with the siRNA as optimized before and 72h post-transfection viability and caspase 3/7 was measure by luminescence using CellTiter-Glo Luminiscent Cell Viability Assay and Caspase-Glo 3/7 Assay (Promega). For each cell line and gene, we obtained a value for viability after siRNA transfection and a value for caspase 3/7. If siRNA transfection decrease viability, we would expect an increase in caspase activity. For every gene and cell line, we looked at the ratio (Caspase/Viability). We select to move forward those genes whose ratio (Caspase/Viability)>1.5 in at least 2 of the tumor cell lines assayed. Fourteen genes met our criteria: DUT, FKBP3, PARP1, PGM1, PP2R1A, RAN, A100A6, SART3, SEPT7, SLC35F6, TBCB, WARS, WDR1 and YWHAE.

We have done a literature review to find if those genes have been previously associated with lung cancer or other type of cancer (Table 3). Most of the genes have been described to play a role in cancer. For example: RAN is a potent oncogene that promotes epithelial to mesenchymal transition (EMT) and invasion in NSCLC,[6] SART3 is a SCC antigen capable of inducing CD8+ T cell activation,[7] FKBP3 promotes NSCLC proliferation and overexpression in NSCLC correlates with poor survival,[8] PARP1 overexpression in lung cancer correlate with poorer survival and enhances AC metastasis,[9, 10] and S100A6 is a diagnostic marker found in serum of patients with stage I/II NSCLC.[11] Based on the literature search, we discarded those genes whose expression is associated with improved survival or decreased tumor growth in some cancers, that is they can act as tumor suppressors (SEPT7, WDR1, YWHAE and PPP2R1A), and we prioritized to move forward those genes that have shown to play a role in lung cancer. We selected 5 genes to determine immunogenicity and for Th1 epitope screening: FKBP3, PARP1, RAN, S100A6 and SART3.

Subtask 1.3. Months 5-7. Determine whether proteins overexpressed in lung cancer cell lines and of functional relevance encode proteins which are immunogenic.

1.3.A. Evaluate de-identified patient samples as compared to controls for autoantibodies to the candidate proteins by ELISA

This sub-aim is currently ongoing.

We have obtained plasma samples from the Seattle Cancer Care Alliance (SCCA) biorepository. Samples include 51 lung cancer patients and 25 normal donors. Our criteria for selection was: male, with NSCLC, early stage (I/II), smokers or former smokers with 15+ years of smoking history. The normal donors were: male, without lung cancer, without any other type of cancer or respiratory diseases, and non-smokers. Summary of the samples characteristics is shown in Table 4.

1.3.B. Validate ELISA responses by western blot.

Work has not yet begun on this sub-aim.

TASK 2: Determine epitopes derived from the identified antigens that will preferentially elicit a Th1 CD4+ T cell response

Subtask 2.1. Months 7-8. Identify promiscuous high affinity binding class II epitopes derived from lung cancer prevention antigens (i.e. universal epitopes).

2.1.A. Use web-based algorithms to analyze candidate protein sequences

Studies by our group have demonstrated that those class II epitopes, derived from self-proteins and most likely to induce a Th response, are epitopes that bind with high affinity across multiple HLA-class II alleles. I have used available modeling programs (SYFPEITHI, Propred, Rankpep) to predict binding affinity. I focused on a panel of 14 HLA-class II alleles which target >90% of major ethnic groups. Our group has developed a multi-algorithm approach for prediction of class II epitopes that are presumed high affinity binders to multiple alleles. After the algorithm, each aminoacid in the protein receives a score number relative to the number of HLA-class II alleles that can bind and the strength of the binding. The score for each aminoacid position on the protein can be shown as a graph or as a HEATMAP (Figure 2). Regions with highest scores have a highest probability of elicit the CD4+ T cell response that we need for our vaccines. Each of the 5 genes selected from the proteomic and siRNA analysis have been screened in silico to identify the epitopes derived from those proteins that will preferentially elicit a Th response. As an example, the score graph for RAN is shown. Figure 2 shows the score graph for RAN with the HEATMAP of one of the peptides selected for analysis.

2.1.B. Select extended epitopes spanning high affinity regions of sequences, discard those with <80% homology to mouse

Table 5 summarizes the sequences of the peptides identified for each protein and the percentage of homology with the corresponding murine protein. Since the validation of the peptide vaccines will be analyzed in mice, we looked for peptides highly homologous between human and mouse. For each protein, we have selected those peptides with higher scores and >80% homology with the mouse protein. A total of 30 peptides were selected for analysis, 6 for RAN, 7 for FKBP3, 6 for SART3, 4 for S100A6 and 7 for PARP1.

2.1.C. Synthesize corresponding peptides.

This aim is currently ongoing. Peptides have been ordered.

Subtask 2.2. Months 9-12. Identify class II epitopes that preferentially induce antigen specific Th1 (IFN-g) rather than Th2 (IL-10).

Work has not yet begun on this sub-aim.

2.2.A. Screen sample PBMC via interferon-gamma ELISPOT for each candidate antigen.

2.2.B. Screen sample PBMC via IL-10 ELISPOT for each candidate antigen.

2.2.C. Identify those peptides with the greatest incidence and magnitude of antigen specific interferon-gamma secreting T cells as compared to IL-10 secreting T cells. Selection of those peptides that have low immune suppressive potential (low IL-10 response) as candidates for a vaccine.

TASK 3: Evaluate the antigen specific Th1 vaccines for prophylactic efficacy in mice models of lung cancer.

Work has not yet begun on this Task.

Subtask 3.1. Months 15-17. Evaluate the immunogenicity of candidate epitopes.

3.1.A. Construct individual vaccine peptides based on Task 2 and immunize mice. Vaccinate with 100 ug of each individual peptide, 3 times at 10-14 day intervals

3.1.B. Determine the immunogenicity of each antigen via interferon-gamma ELISPOTs for both the individual peptides as well as corresponding recombinant protein.

Subtask 3.2. Months 17-24. Evaluate the prophylactic efficacy of vaccination.

3.2.A. Immunize mice with a multi-antigen peptide vaccine and start treatment with NTCU to induce lung cancer to the vaccinated animals

3.2.B. Image mice by MRI at 10 and 18 weeks of NTCU treatment to follow up the development of lung lesions

3.2.C. Evaluate the efficacy of vaccination at 20 weeks by histological analysis of lungs.

3.2.D. Evaluate the immune responses to vaccine peptides and corresponding proteins in splenocytes by ELISPOT assays.

4. KEY RESEARCH ACCOMPLISHMENTS:

- We have identified 154 proteins overexpressed in lung cancer cell lines, as potential targets for a lung cancer vaccine.
- 14 genes, coding identified overexpressed proteins in lung cancer, were found critical for proliferation and survival of the lung cancer cell lines by siRNA screening. For all of them there is published data describing roles in cancer.
- We have mapped HLA-DR epitopes for 5 of the proteins identified.

5. CONCLUSION:

Overexpressed protein in tumors can be immunogenic and are good candidates for cancer vaccines. We have been able to identify proteins overexpressed in NSCLC cell lines that are critical for tumor cell survival and proliferation. These proteins, overexpressed and relevant for tumor growth, are good targets for anti-tumor immunotherapies. Future plans:

- Evaluation of the antigens for immunogenicity in patients with lung cancer. We will study the presence antibodies to the identified antigens in plasma samples from lung cancer patients.
- Screening of the HLA-DR epitopes to identify those capable of inducing a Th1 immune response. ELISPOT screening for IFN-g (Th1) and IL-10 (Th2) will identify Th1 epitopes to be included in vaccines as well as those Th2 epitopes that should be excluded.
- Evaluate the efficacy of a multi-antigen Th1 vaccine in mouse models of lung cancer.

6. PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:

Nothing to report

7. INVENTIONS, PATENTS AND LICENSES:

Nothing to report

8. REPORTABLE OUTCOMES:

Nothing to report

9. OTHER ACHIEVEMENTS:

Nothing to report

10. TRAINING:

As part of my training, I had one-on-one meetings with my primary mentor, Dr. Nora Disis, every two weeks to assess the progress of the project, discuss and review data, and help with problem solving/future plans. In addition, I have met with my co-mentor, Dr. Renato Martins, quarterly to evaluate progress and make sure the experiments are done accordingly to the real necessities of the patients to facilitate future translation of the pre-clinical data to the clinic. I have participated in the weekly meetings in the Tumor Vaccine Group and the Clinical Meetings in Dr. Martins group. I have also attended to UW immunology Lectures and the seminars offered by the Institute for Translational Health Sciences in the area of biostatistics, grant writing and clinical trials. I have attended the following modules within the Summer Institute in Bioinformatics Course at the University of Washington: Introduction to R, Integrative genomics, GWAS and Sequencing Data, and Pathways and network analysis.

11. COLLABORATIVE AWARDS:

Not applicable

12. QUAD CHARTS

Not applicable

13. REFERENCES:

1. Spira, A., et al., *Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer.* Nat Med, 2007. **13**(3): p. 361-6.
2. Woodard, K.M. and C.J. Chapman, *Lung cancer - can autoantibodies provide an aid to diagnosis?* Expert Opin Med Diagn, 2008. **2**(8): p. 911-23.
3. Lowe, F.J., et al., *A novel autoantibody test for the detection of pre-neoplastic lung lesions.* Mol Cancer, 2014. **13**: p. 78.
4. Goodell, V., et al., *Level of HER-2/neu protein expression in breast cancer may affect the development of endogenous HER-2/neu-specific immunity.* Mol Cancer Ther, 2008. **7**(3): p. 449-54.
5. Disis, M.L., *Immune regulation of cancer.* J Clin Oncol, 2010. **28**(29): p. 4531-8.
6. Ning, J., et al., *Ran GTPase induces EMT and enhances invasion in non-small cell lung cancer cells through activation of PI3K-AKT pathway.* Oncol Res, 2013. **21**(2): p. 67-72.
7. Suefuji, Y., et al., *Expression of SART3 antigen and induction of CTLs by SART3-derived peptides in breast cancer patients.* Br J Cancer, 2001. **84**(7): p. 915-9.
8. Zhu, W., et al., *FKBP3 Promotes Proliferation of Non-Small Cell Lung Cancer Cells through Regulating Sp1/HDAC2/p27.* Theranostics, 2017. **7**(12): p. 3078-3089.
9. Choi, E.B., et al., *PARP1 enhances lung adenocarcinoma metastasis by novel mechanisms independent of DNA repair.* Oncogene, 2016. **35**(35): p. 4569-79.
10. Xie, K.J., et al., *Expression of ERCC1, MSH2 and PARP1 in non-small cell lung cancer and prognostic value in patients treated with platinum-based chemotherapy.* Asian Pac J Cancer Prev, 2014. **15**(6): p. 2591-6.
11. Wang, T., et al., *Diagnostic significance of S100A2 and S100A6 levels in sera of patients with non-small cell lung cancer.* Tumour Biol, 2016. **37**(2): p. 2299-304.

14. APPENDICES: Figures and Tables

Cell Line	Type	Stage	iTRAQ Isobaric Tag
Nuli-1	Normal	na	113, 121
BEAS-2B	Normal	na	114
HCC1588	SCC	I	115
HCC1313	SCC	IIIa	116
H520	SCC	not known	119
H522	AC	II	117
HCC461	AC	IIIa	118

Table 1. Cell lines used for quantitative mass-spectroscopy

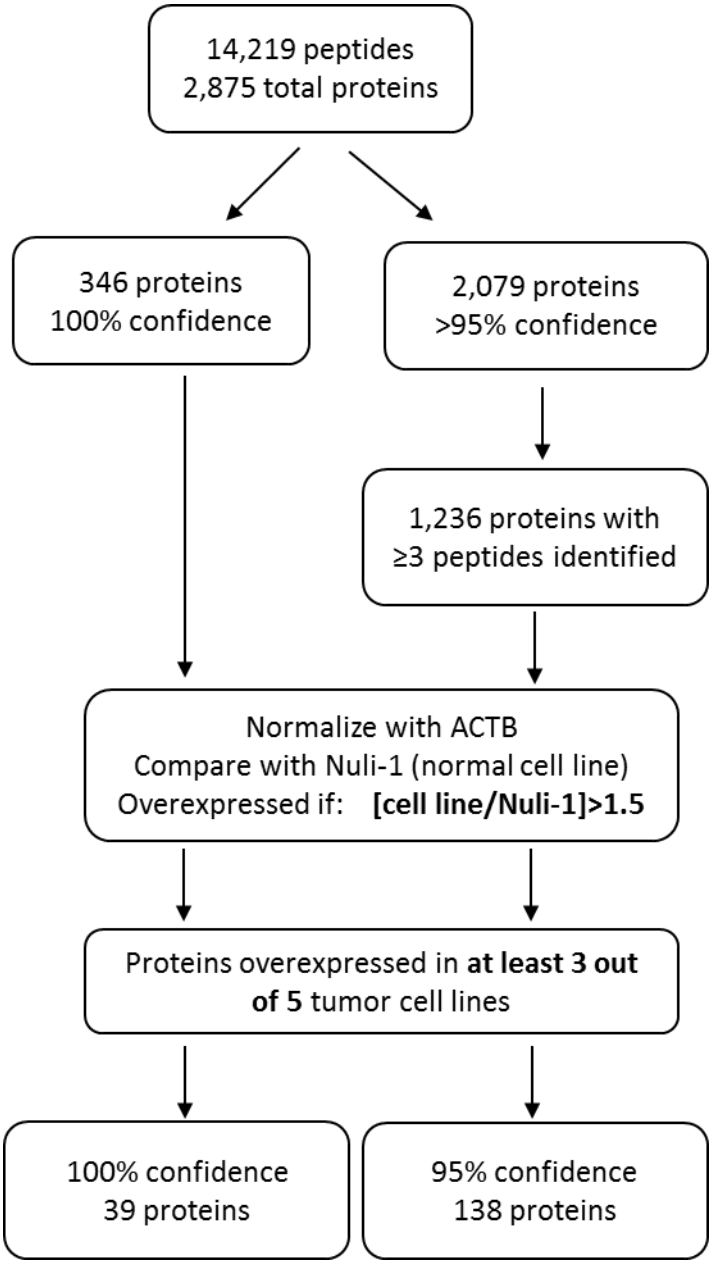

Figure 1. Rationale used to identify proteins overexpressed in the tumor versus the normal cell line Nuli-1

Cell line	GENE	Test Mean	Population Mean	Test/Population	pValue Pairwise
H520	AHCY	48.1	66.6	0.72	0.0130
H520	CYC1	43.5	82.3	0.53	0.0004
H520	GPI	44.1	68.1	0.65	0.0002
H520	PARP1	48.7	69.1	0.71	0.0022
H520	UCHL1	43.9	60.5	0.73	0.0022
H520	WDR1	44.5	64.2	0.69	0.0001
H520	YWHAE	48.7	68.3	0.71	0.0004
H522	CACYBP	50.1	67.4	0.74	0.0008
H522	CES1	56.7	77.6	0.73	0.0003
H522	DUT	44.3	62.2	0.71	0.0139
H522	FKBP3	38.4	53.4	0.72	0.0056
H522	PPP2R1A	38.3	57.4	0.67	0.0070
H522	S100A6	41.5	70.7	0.59	0.0005
H522	SART3	45.3	70.8	0.64	0.0001
HCC1588	PGM1	33.8	55.6	0.61	0.0002
HCC1588	RAN	22.3	32.8	0.68	0.0056
HCC1588	7-Sep	50.0	71.1	0.70	0.0001
HCC1588	SLC35F6	32.5	57.4	0.57	0.0004
HCC1588	TBCB	42.4	65.7	0.65	0.0001
HCC1588	WARS	31.4	52.0	0.60	0.0008
HCC461	FUS	56.5	77.7	0.73	0.0017

Table 2. siRNA viability screening. Genes for which [(mean of viability tumor cell line) / (mean of viability in the normal cell line)] <0.75 with a p-value <0.1 for any of the cell lines assayed.

GENE	Mouse homology	Full Name	Cancer Involvement
DUT	81.5	Deoxyuridine Triphosphatase	• Overexpression correlates with poor prognosis and shorter DFS in hepatocellular carcinoma • High levels of DUT in colon cancer predict tumor resistance to chemo, shorter time to progression, and shorter OS
FKBP3	92.1	FK506 Binding Protein 3	• Overexpression in NSCLC correlates with poor survival. Promotes NSCLC proliferation through regulation of SP1/HDAC2/p27
PARP1	86.5	Poly(ADP-Ribose) Polymerase 1	• Overexpressed in 55.9% NSCLC and correlate with decreased survival time • Enhances lung adenocarcinoma metastasis • Polymorphisms can contribute to lung cancer susceptibility • Proteomic profiling identified PARP1 as target for small cell lung cancer
PGM1	85.4	Phosphoglucomutase 1	• Involved in glioma and uterine cancer development • Role in cancer cell survival under hypoxia conditions
PPP2R1A	90.6	Protein Phosphatase 2 Scaffold Subunit A alpha	• Mutations in PPP2R1A are associated with development of endometrial, ovarian and uterine cancer • Mutations in this gene act as dominant negative and promote tumor cell growth
RAN	93.0	RAN, Member RAS Oncogene Family	• Oncogene, promote cancer progression. Induces EMT and invasion in NSCLC
S100A6	93.3	S100 Calcium Binding Protein A6	• Increased cytoplasmic expression is associated with lung adenocarcinoma progression • S100A6 is found in serum of stage I/II NSCLC patients - potential biomarker • Binds ANXA2 and promote cancer cell motility in pancreatic and colon cancer • Overexpression enhances tumorigenesis and is associated with poor prognosis in gastric cancer, renal cell carcinoma, hepatocellular carcinoma. Also associated with breast, osteosarcoma, pancreas and ovarian cancer.
SART3	83.3	Squamous Cell Carcinoma Antigen Recognized By T-Cells 3	• Antigen for CD8 CTL activation in breast, colon, prostate, renal, gastric and oral cancer • Vaccines evaluating SART3 HLA-I epitopes showed no adverse effects in patients and increased immunity against colon cancer
SEPT7	95.0	Septin 7	• Upregulation and expression is required for growth of hepatocellular carcinoma, and for migration and invasion of human breast cancer • Upregulation in glioma inhibits invasion
SLC35F6	85.8	Solute Carrier Family 35 Member F6	• Unfavorable prognostic marker in lung cancer • Expressed in multiple cancer, including breast, endometrium, liver, lung, melanoma, renal, among others.
TBCB	85.9	Tubulin Folding Cofactor B	• no association with cancer reported
WARS	87.1	Tryptophanyl-TRNA Synthetase	• Overexpression enhances oral cancer cell invasiveness • Spicing forms influences pancreatic cancer metastasis • Play a role in colorectal and cervical cancer
WDR1	87.7	WD Repeat Domain 1	• Relevant in breast cancer invasiveness • Overexpression in glioma is associated with poor prognosis • In ovarian carcinoma resistant to chemotherapy, under expression of WDR1 is associated with poor overall survival
YWHAE	93.5	Tryptophan 5-Monooxygenase Activation Protein Epsilon, 14-3-3 epsilon	• Role in endometrial stromal sarcoma- fusion protein by rearrangement of YWHAE-FAM22 promote tumor growth • In gastric cancer, downregulation of the gene induces proliferation, invasion and migration of cancer cells

Table 3. List of genes for which the Ratio (Caspase/Viability)>1.5 in at least 2 of the tumor cell lines assayed. Percentage of homology with mouse orthologue protein and previously described role in cancer is shown.

	NSCLC	Control
Sample number	51	25
Age	70 (55-88)	51 (22-69)
Stage		
Ia	25 (49%)	-
Ib	13 (25%)	-
IIa	8 (16%)	-
IIb	4 (8%)	-
Smoker		
Current	15 (29%)	-
Former	35 (69%)	-
Years smoked	38.5 (10-60)	-
Pack years smoked*	49.4 (6-180)	-

Table 4. Human subjects characteristics. Plasma samples from early stage NSCLC patients and normal donors. Age, years smoked and pack years are shown as "average (range min-max)".
* Pack years smoked is define as the number of packs of cigarettes smoked per day multiplied by the number of years the person has smoked.

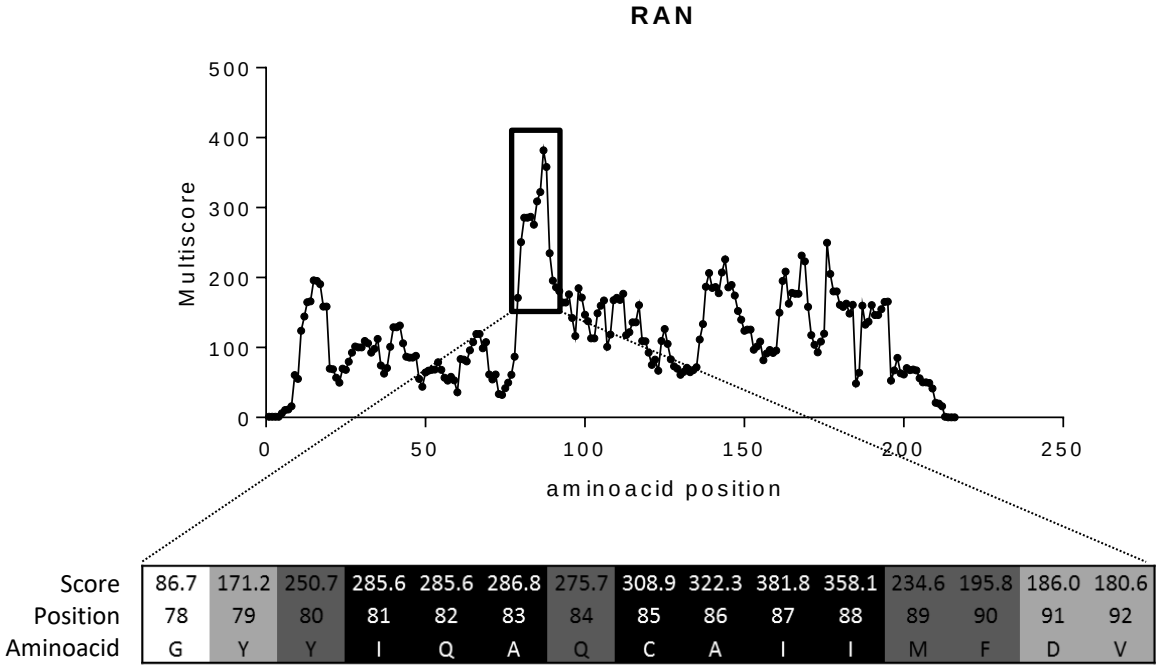

Figure 2. Identification of Th epitopes in RAN. Multi-score graph shows epitope binding scores for every aminoacid along the RAN protein. RAN p78-92 is shown as an example of a peptide spanning 15 aminoacid containing high affinity binding epitopes across multiple HLA-II alleles. Binding scores are assigned by color-coded quartiles. Top 25% scores in black, 50-75% dark grey, 25-50% light grey and 0-25% in white. Darker areas are considered "immunologic hot-spots."

Protein name, ID#, aa	Peptide #	Sequence	Sum Score	Mouse homology
RAN	p9-23	VQFKLVLVGDGGTGK	1859	100%
P62826.3	p78-92	GYYIQAQCAIIMFDV	3810	100%
216 aa	p93-107	TSRVTYKNVPNWHRD	2208	100%
	p136-150	IVFHRKKNLQYYDIS	2587	100%
	p151-165	AKSNYNFEKPFLWLA	1910	100%
	p166-180	RKLIGDPNLEFVAMP	2485	100%
FKBP3	p19-33	QLPKKDIIKFLQEH	1954	93%
Q00688.1	p34-50	SDSFLAEHKLLGNIKNV	2995	100%
244 aa	p49-63	NVAKTANKDHLVTAY	2034	93%
	p62-76	AYNHLFETKRFKGTE	2219	93%
	p130-144	VVHCWYTGTLQDGT	1832	93%
	p161-175	PLSFKVGVGKVIRGW	2265	100%
	p176-190	DEALLTMSKGEKARL	2117	100%
SART3	p229-243	EAYREFESAIVEAAR	917	100%
Q15020.1	p244-260	LEKVHSLFRRQLAIPLY	1591	100%
963 aa	p422-436	QAYLDYLRRRVDFKQ	883	87%
	p441-455	ELEELRAAFTRALEY	809	87%
	p730-744	GEVVQIRPIFSNRGD	1184	100%
	p791-805	FRYSTSLEKHKLFI	791	93%
S100A6	p9-24	IGLLVAIFHKYSGREG	5275	100%
P06703.1	p35-49	KELIQKELTIGSKLQ	2974	100%
90 aa	p49-63	QDAEIARLMEDLDRN	2678	93%
	p70-87	FQEYVTFLGALALIYNEA	6400	94%
PARP1	p32-46	SLRMAIMVQSPMFDG	620	100%
P09874.4	p51-65	WYHGSCFWKVGHSIR	853	87%
1041 aa	p582-596	RYWIFRSWGRVGTVI	1206	93%
	p851-866	FKQLHNRRLLWHGSRT	1067	94%
	p865-879	RTTNFAGILSQGLRI	787	100%
	p893-907	KGIYFADMVSKSANY	820	100%
	p990-1005	IVYDIAQVNKYLLKL	1019	94%

Table 5. Peptides selected for ELISPOT IFN-g (Th1) and IL-10 (Th2) screening.